

REPORT 7336  
SEPTEMBER 1973

# Operator Factorization and Other Aspects of the Analysis of Linear Systems

PER HAGANIDER

**LTH**

Division of Automatic Control · Lund Institute of Technology





OPERATOR FACTORIZATION AND OTHER ASPECTS  
OF THE ANALYSIS OF LINEAR SYSTEMS

av

Per Hagander  
tekn lic, Ssk

Akademisk avhandling som för avläggande av teknisk doktorsexamen vid tekniska fakulteten vid universitetet i Lund kommer att offentligen försvaras i sal M:A, Maskinhuset, Lunds Tekniska Högskola, fredagen den 7 december 1973 kl. 10.15.

To Barbro



AV Centralen  
Lund 1973

## OPERATOR FACTORIZATION AND OTHER ASPECTS OF THE ANALYSIS OF LINEAR SYSTEMS.

Applications of control theory appear in many branches of engineering, biology and economy. The basic questions are the same, although different aspects may be emphasized. Control theory discusses how to manipulate a process and how to describe the resulting behaviour. A system is called static if the current behaviour depends only on the current values of the control signals. For dynamical systems it may depend also on old values. Causality implies that future control actions must not have any influence on the present behaviour.

The aim of control is to use the available input signals to make the system behave as well as possible in spite of uncertainty and disturbances. A regulator determines the input signals from measurements of the process.

The formulation of mathematical models for a system usually involves many simplifications made in order to reduce complexity. Linearization is frequently used. Uncertainty can be taken into account by using random variables. The objective of the system is also formalized in mathematical terms.

### Models for Linear Dynamical Systems.

There are mainly two types of descriptions of dynamical systems. If only the input-output behaviour is modelled it is natural to talk about an external description. The output at time  $t$  is then expressed as a weighted sum of the values of the input at past times, i.e.

$$y(t) = \int_{-\infty}^t h(t,s)u(s)ds \quad (1)$$

for continuous time and

$$y(t) = \sum_{-\infty}^t h(t,s)u(s) \quad (2)$$

for discrete time. The function  $h$  is called the weighting function.

For many processes there exists much knowledge of the internal behaviour. Physical laws may govern the dynamics, and it may be possible to define a number of variables, the state, that summarizes the influence of the past. The concept originates from classical mechanics, where positions and velocities are typical state variables.

First order linear differential or difference equations describing the change of the state are natural models. The measurements are assumed to be linear combinations of the state variables and the inputs. Vector and matrix notation is widely used. A continuous time system could thus be written

$$\begin{cases} \frac{d}{dt} x(t) = Ax(t) + Bu(t) & x(t_0) = x_0 \\ y(t) = Cx(t) + Du(t) \end{cases} \quad (3)$$

and correspondingly in the discrete time case:

$$\begin{cases} x(t+1) = \phi x(t) + \Gamma u(t) & x(t_0) = x_0 \\ y(t) = \theta x(t) + Du(t) \end{cases} \quad (4)$$

It is possible to obtain an input-output description

from all internal descriptions like (3) and (4). In fact, the models (1) and (2) are somewhat more general since it may not be possible to find a finite number of state variables.

The external model is just a linear map, an operator, that relates output signals to the input signals. If there is also an internal state variable description, it further specifies the operator.

### Two Problems.

Having some noisy measurements it is a natural problem to try to describe the behaviour of a system, for instance by estimating the state. A good estimate should be close to the true value. It is therefore reasonable that the estimation error should have zero mean value and minimal variance. Another problem is to find inputs to drive a system back to a desired state, for instance zero, after some disturbance. For many linear systems there is an input signal, that makes the state zero in a very short time at the expense of control energy. A measure is usually defined that balances the deviation at some final time,  $t_1$ , against the required control energy. The behaviour during the trajectory may also be taken into account. The most common loss function is the quadratic form

$$J = x^T(t_1)Q_0x(t_1) + \int_{t_0}^{t_1} x^T(t)Q_1x(t)dt + \\ + \int_{t_0}^{t_1} u^T(t)Q_2u(t)dt \quad (5)$$



for continuous time and for discrete time

$$J = \int_{t=t_0}^{t_1} x^T(t) Q_1 x(t) + \int_{t=t_0}^{t_1} u^T(t) Q_2 u(t) \quad (6)$$

where  $Q_0$ ,  $Q_1$  and  $Q_2$  are symmetric and positive semi-definite. In order to be able to penalize infinite values of the inputs it is usually assumed that  $Q_2 > 0$ .

#### Methods for Solution.

Around 1940 Wiener [11] and Kolmogorov [9] solved these two problems using the weighting function representation. Wiener used calculus of variations. He found that the problem could be reduced to the solution of an integral equation, the Wiener-Hopf equation. Wiener solved the equation by function theoretic methods. In simple cases this reduces to the problem of factorizing a polynomial into a part with zeroes in the left half-plane and another part with zeroes in the right half-plane only (spectral factorization). Twenty years later Bellman [2] and Kalman [6] etc. applied dynamic programming and the Hamilton-Jacobi equation to the state variable representation and found a formally and computationally different approach. The problem reduced to an initial value problem for an ordinary nonlinear differential equation, the Riccati equation.

In this thesis the causal linear operator representation is used to formulate the two linear problems discussed above. The solution is obtained by factorizing certain operators in order to solve operator equations corresponding to the Wiener-Hopf equation. The state variable



description is utilized, and the Riccati equation gives the factorization directly in the time domain.

The thesis is composed of the following publications:

- I. The Use of Operator Factorization for Linear Control and Estimation. Automatica 9 pp 623-31 (1973).
- II. Inversion of a Dynamical System by an Operator Identity. Automatica 8 pp 361-362 (1972).
- III. A New Proof and an Adjoint Filter Interpretation for Linear Discrete Time Smoothing. Report 7330, Lund Institute of Technology, Division of Automatic Control.
- IV. Kalman Filters for Processes with Unknown Initial Values. Report 7332, Lund Institute of Technology, Division of Automatic Control.
- V. Numerical Solution of  $A^T S + SA + Q = 0$ . Information Sciences 4 pp 35-50 (1972).

The first three parts are devoted to the operator approach to the control and estimation problems. Part I treats the continuous time problems, part II deals with the inversion of dynamical systems and part III demonstrates the operator technique for discrete time applied to the state estimation problem.

One problem of the linear theory that does not seem to have a satisfactory solution is estimation for processes with unknown initial state. This problem is discussed in part IV.

The application of the linear theory to any realistic

problem requires substantial numerical computation. Since the different approaches to the linear problem also lead to different computational methods it would be interesting to analyse the numerical aspects. No complete study has been made. Part V discusses a simplified problem, which contains many of the more general aspects.

### Operator Notation.

Estimation and optimization problems are easy to solve for static systems. Dynamical systems can in fact also be considered as static systems in spaces of much larger dimensions. The discrete time systems on finite intervals, say  $[0, t_1]$ , can equivalently be analyzed using vectors like

$$\underline{x} = \begin{bmatrix} x(0) \\ x(t) \\ \vdots \\ x(t_1) \end{bmatrix}$$

The system (4) in weighting function form is:

$$\begin{cases} x(t) = \phi(t, 0)x_0 + \sum_{s=0}^{t-1} \phi(t, s+1)r(s)u(s) \\ y(t) = \theta x(t) + Du(t) \end{cases}$$

It can also be written as

$$\begin{cases} \underline{x} = \underline{g}x_0 + \underline{L}ru \\ \underline{y} = \underline{\theta}\underline{x} + \underline{D}u \end{cases} \quad (7)$$

using the block matrices:

$$\underline{L} = \begin{bmatrix} 0 & & \\ I & & 0 \\ \phi(t_1, 1) & I & 0 \end{bmatrix}$$

$$\underline{r} = \begin{bmatrix} r(0) & & \\ & & 0 \\ 0 & & r(t_1) \end{bmatrix}$$

$$\underline{\theta} = \begin{bmatrix} \theta(0) & & \\ & & 0 \\ 0 & & \theta(t_1) \end{bmatrix}$$

$$\underline{g} = \begin{bmatrix} I \\ \phi(1, 0) \\ \vdots \\ \phi(t_1, 0) \end{bmatrix}$$

$$\underline{D} = \begin{bmatrix} D(0) & & \\ & & 0 \\ 0 & & D(t_1) \end{bmatrix}$$

The structure of the matrices should be noted. The matrix  $\underline{L}$  is lower block triangular, a consequence of causality. For time constant  $\phi$  it also has a band structure. The matrices  $\underline{r}$ ,  $\underline{\theta}$  and  $\underline{D}$  are block diagonal. The signals  $\underline{u}$ ,  $\underline{x}$  and  $\underline{y}$  could be considered to be functions on the discrete time interval  $[0, t_1]$  instead of being vectors. The block matrices then correspond to linear operators.

Introducing appropriate block diagonal matrices  $\underline{Q}_1$  and  $\underline{Q}_2$  the loss function (6) can be written as

$$J = \underline{x}^T \underline{Q}_1 \underline{x} + \underline{u}^T \underline{Q}_2 \underline{u} \quad (8)$$

If the expression for  $\underline{x}$  is inserted into (8) then

$$J = \underline{u}^T (\underline{Q}_2 + \underline{r}^T \underline{L}^T \underline{Q}_1 \underline{L} \underline{r}) \underline{u} + 2 \underline{u}^T \underline{r}^T \underline{L}^T \underline{Q}_1 \underline{g} x_0 + x_0^T \underline{g}^T \underline{Q}_1 \underline{g} x_0 \quad (9)$$

The dynamic optimization problem is thus rewritten as a static problem. The solution is easily obtained as



$$\underline{u} = - (\underline{Q}_2 + \underline{\Gamma}^T \underline{L}^T \underline{Q}_1 \underline{L} \underline{\Gamma})^{-1} \underline{\Gamma}^T \underline{L}^T \underline{Q}_1 \underline{g} x_0 \quad (10)$$

provided that the inverse exists.

For the estimation problem let  $\underline{u} = 0$  in (7) and add white noise terms  $\underline{v}$  and  $\underline{e}$  with block diagonal covariance matrices  $\underline{R}_1$  and  $\underline{R}_2$ ,  $\underline{R}_2 > 0$ . Assume also that  $x_0$  is a random vector with covariance matrix  $\underline{R}_0$ . Thus

$$\begin{cases} \underline{x} = \underline{g} x_0 + \underline{L} \underline{v} \\ \underline{y} = \underline{\theta} \underline{x} + \underline{e} \end{cases} \quad (11)$$

which is also a static problem. Assume for simplicity that all mean values are zero. The minimal variance linear estimate  $\hat{\underline{x}}$  of  $\underline{x}$  based on  $\underline{y}$ , i.e.  $y(t)$ ,  $t \in [0, t_1]$ , can be expressed in terms of covariance matrices by the projection theorem

$$\hat{\underline{x}} = \underline{R}_{\underline{x}\underline{y}} \underline{R}_{\underline{y}}^{-1} \underline{y} = \underline{R}_{\underline{x}} \underline{\theta}^T (\underline{\theta} \underline{R}_{\underline{x}} \underline{\theta}^T + \underline{R}_2)^{-1} \underline{y} \quad (12)$$

provided that the matrix  $\underline{R}_{\underline{y}}$  is invertible.

Both optimization and estimation are thus solved as static problems in larger spaces. In fact, many problems in linear systems theory can be explained easily in this setting. Typical examples are controllability, observability and invertibility.

The continuous time problems are conceptually more difficult. Finite dimensional spaces are no longer sufficient. Vectors indexed by an interval of the real line could be used, but the function space interpretation is the most natural. The vector - matrix analogy is useful for intuition, so one can speak about triangular and diagonal operators.

The continuous time system description is introduced in part I. In analogy with (7) the system (3) is written

$$\begin{cases} \underline{\dot{x}} = \underline{g}x_0 + \underline{L}Bu \\ \underline{y} = \underline{C}x + \underline{D}u \end{cases}$$

It is shown that the adjoints  $\underline{L}^*$  and  $\underline{g}^*$  can be described by the dual system, and that the operator  $\underline{L}$  is only left invertible. The analogues of (10) and (12) are also shown. Systems with nonsingular  $\underline{D}$  are invertible dynamical systems, and in part II it is shown how the inverse can be obtained using operators. Let  $x_0$  be zero, then the inverse system is

$$\underline{u} = (\underline{C}\underline{L}\underline{B} + \underline{D})^{-1} \underline{y} = (\underline{D}^{-1} - \underline{D}^{-1} \underline{C} \underline{M} \underline{B} \underline{D}^{-1}) \underline{y} \quad (13)$$

$$\underline{x} = \underline{L} \underline{B} \underline{u} = \underline{L} \underline{B} (\underline{C} \underline{L} \underline{B} + \underline{D})^{-1} \underline{y} = \underline{M} \underline{B} \underline{D}^{-1} \underline{y} \quad (14)$$

where  $\underline{M}$  is a new triangular operator. The result corresponds to an elementary matrix lemma. In parts I and III frequent use is made of this inversion lemma.

The advantage of reformulating the problems as static problems is the simple formal solutions (10) and (12). The drawback is that the recursive, dynamic nature of the solutions is lost. A fundamental idea of this thesis is to use the static formulation with its simple formal solution and then provide methods by which recursive equations for the solution can be obtained.

### Solution Using Operator Factorization.

Both the estimation and the optimization problems require the solution of a linear equation (12) or (10). The usual way of solving such equations in finite spaces is by Gauss elimination, i.e. triangularization and recursive solution of a backward and a forward system.

In the discrete time the number of operations required for Gauss elimination would be of the order of  $[n(t_1-t_0)]^3$ . The structure of the two operators to invert in (12) and (10) for the estimation and control problems can, however, be exploited to reduce the computations. For example if  $\underline{R}_1 = \underline{I}$ ,  $\underline{R}_2 = \underline{I}$ ,  $\underline{\Theta} = \underline{I}$ ,  $\underline{C} = \underline{I}$ ,  $\underline{R}_0 = 0$ , then  $\underline{R}_Y$  is given by

$$\underline{R}_Y = \underline{I} + \underline{L}\underline{L}^* \quad (15)$$

both in the discrete and continuous time cases. A symmetric factorization of  $\underline{R}_Y$  into lower and upper triangular operators can be obtained using the Riccati equation.  $\underline{R}_Y$  can be written

$$\underline{R}_Y = \underline{I} + \underline{L}\underline{L}^* = (\underline{I} + \underline{L}\underline{P})(\underline{I} + \underline{P}\underline{L}^*) \quad (16)$$

in the continuous time case and for discrete time

$$\underline{R}_Y = \underline{I} + \underline{L}\underline{L}^* = (\underline{I} + \underline{P} + \underline{L}\Phi\underline{P})(\underline{I} + \underline{P})^{-1}(\underline{I} + \underline{P} + \underline{P}\Phi^T\underline{L}^T) \quad (17)$$

provided that  $\underline{P}$  is the diagonal operator with  $P(t)$  satisfying the continuous or discrete time Riccati equation with zero initial value. These identities and their generalizations are shown in parts I and III.

The continuous time solution in part I will now be demonstrated. Using  $\underline{P}$  the operator  $\underline{R}_{xy}$  is

$$\underline{R}_{xy} = \underline{L}\underline{L}^* = \underline{P}\underline{L}^* + \underline{L}\underline{P}(\underline{I} + \underline{P}\underline{L}^*)$$

so that

$$R_{\underline{X}\underline{Y}} R_{\underline{Y}}^{-1} = \underline{L}\underline{L}^*(\underline{I} + \underline{L}\underline{L}^*)^{-1} = \underline{P}\underline{L}^*(\underline{I} + \underline{P}\underline{L}^*)^{-1}(\underline{I} + \underline{L}\underline{P})^{-1} + \underline{L}\underline{P}(\underline{I} + \underline{L}\underline{P})^{-1}$$

Inversion of dynamical systems gives

$$(\underline{I} + \underline{L}\underline{P})^{-1} = \underline{I} - \underline{M}\underline{P}, \quad \underline{L}\underline{P}(\underline{I} + \underline{L}\underline{P})^{-1} = \underline{M}\underline{P}, \quad \underline{L}^*(\underline{I} + \underline{P}\underline{L}^*)^{-1} = \underline{M}^*$$

where  $\underline{M}$  is a new lower triangular operator like  $\underline{L}$ . Thus

$$R_{\underline{X}\underline{Y}} R_{\underline{Y}}^{-1} = (\underline{L}\underline{P} + \underline{P}\underline{M}^*)(\underline{I} - \underline{M}\underline{P}) = \underline{M}\underline{P} + \underline{P}\underline{M}^*(\underline{I} - \underline{M}\underline{P}) \quad (18)$$

gives the so called smoothing estimate. The discrete time estimate is given in part III. Eq. (18) represents causal and anticausal integral operators. Using forward and backward initial value differential equations it can be reformulated as by Bryson and Frazier [3]. It is also interesting that the operator  $\underline{M}\underline{P}$  gives the filter estimate, i.e. the estimate of  $x(t)$  using  $y(s)$  only for  $s \leq t$ . This corresponds to the forward equation.

Instead of a high dimensional Gauss elimination, a simpler factorization and some dynamical system inversions directly give the formulas used in practice. The matrix Riccati equation with the dimension of the state plays the central role of the factorization. It is thus found how the internal structure, which is lost in the input output formulation, can be exploited to obtain recursive algorithms.

The continuous time regulator problem is treated in part I, including also restricted end point, deterministic disturbance known in advance and disturbances described by stochastic processes.



### Earlier Work.

The operator notation, where the signals are considered as time functions, has been used before, e.g. by Bellman [1] and Kailath [5], but mostly descriptively. The connection between Riccati equations and Fredholm and Wiener-Hopf integral equations was shown by the work by Kalman and Bucy [8] and has been applied primarily in the field of two point boundary value problems by Schumitsky [10] for the factorization of Fredholm resolvents. Gohberg and Krein [4] have given an important theorem on the existence and uniqueness of such factorizations also for asymmetric problems. Kailath [5] has applied the resolvent identity to signal smoothing, i.e. estimation of  $(y - \hat{y})$ , in his innovations approach to the estimation problem. Using the inverse of (16) in this way there are difficulties when  $C \neq I$  or  $R_0 \neq 0$ . Some extensions of the solutions discussed here to infinite dimensional state spaces are possible. The extension to infinite time horizon also requires some caution. Stable systems can be handled without much extra effort. The factorization (16) directly corresponds to spectral factorization of the corresponding Laplace transforms. The system  $I + LP$  is stable and has a stable inverse.

### Processes With Unknown Initial Values.

In the well-established field of linear quadratic control the estimation problem for a system with unknown initial state has not been satisfactorily solved. This problem is treated in part IV. The problem is the dual of optimal control for fixed end point. Intuitively speaking a "dead beat filter" which eliminates the effect of the unknown initial value should be applied. A special case was discussed early by Kalman [7] in connection with observability. In practice one usually solves the ordinary recursive equations starting with a large covariance. This method is numerically ill-conditioned. In part IV the discrete time optimal filter is ob-

tained by letting the initial covariance go to infinity. Two Riccati equations are then obtained, one for the error covariance and one for updating the unobservable subspace. A different type of solution obtained by duality is suggested for the continuous time case.

### Numerical Aspects.

The numerical aspects of the problems discussed above are important for economic realization on digital computers, and by no means fully understood. One example, treated in part V, is the problem of solving  $S$  in the simple  $n \times n$  matrix equation

$$A^T S + SA + Q = 0 \quad (19)$$

This fundamental equation is often called the Lyapunov equation, since it can be used to construct Lyapunov functions. It also arises when evaluating steady state loss functions in linear control and covariance matrices for dynamical systems. More than ten different methods have been proposed emanating both from internal and external system descriptions. Part V gives a survey and classifies them into direct methods, transformation methods and iterative methods. The nine algorithms that seem to be the best ones are coded and tested on a batch of representative  $A$  and  $Q$  matrices. The direct methods which convert (19) to a  $n(n+1)/2$  dimensional linear system thereby losing much structure, are good for small problems, say  $n < 7$ , but unfeasible at present stage for large problems because of large core requirement and long computing time. The iteration methods are the best for larger systems. The required computing time depends also on the numerical condition of the equation. The fastest algorithm is a transformation method but the accuracy is too bad. If the  $A$  matrix is in any special form like Jordan or Companion form or if the corresponding transformations would be of any value in the fur-

ther analysis of a problem, then it should be seriously considered to use a method that takes advantage of this structure.

#### Acknowledgements.

I want to express my sincere gratitude to my adviser Karl Johan Åström, who proposed the operator factorization problems, and who has constantly supported me with ideas and encouragement. I would also like to thank all my colleagues at the Division of Automatic Control, who form a group with an open and stimulating atmosphere. Many improvements have resulted from our frequent discussions. The typing of the manuscripts with all their formulas was done by Gudrun Christensen and Kerstin Palmqvist, and I really appreciate their work. Finally I want to thank my wife Barbro for her understanding and support.

#### References.

- [1] Bellman R.E.: Introduction to the Mathematical Theory of Control Processes, Vol. 1, Academic Press, New York (1967).
- [2] Bellman R.E.: On a Class of Variational Problems, Q. Appl. Math. 14, 353-9 (1957).
- [3] Bryson A.E., Frazier M.: Smoothing for Linear and Nonlinear Dynamic Systems, Aero. Sys. Div., Wright-Patterson AFB, Ohio, Tech. Rept., ASD-TDR-63-119 (1963).
- [4] Gohberg I.C., Krein M.G.: On the Factorization of Operators in Hilbert Spaces, Acta Sci. Math. Szeged. 25, 90-123 (1964).
- [5] Kailath T.: Application of a Resolvent Identity to a Linear Smoothing Problem, SIAM J. Control 7, 68-74 (1969).
- [6] Kalman R.E.: Contributions to the Theory of Optimal Control, Bol. Soc. Mat. Mexicana 5, 102-119 (1960).
- [7] Kalman R.E.: New Methods in Wiener Filtering Theory, in Proc. of 1st Symp. on Eng. Appl. of Random Function Theory and Probability, Bogdanoff J.L., Kozin F. (eds.), Wiley, New York (1963).
- [8] Kalman R.E., Bucy R.S.: New Results in Linear Filtering and Prediction Theory, J. Basic Eng., Trans ASME Ser. D 83, 95-108 (1961).
- [9] Kolmogorov A.N.: Interpolation and Extrapolation of Stationary Random Sequences, Bull. Moscow Univ. USSR, Ser. Math. 5 (1941).
- [10] Schumitsky A.: On the Equivalence between Matrix Riccati Equations and Fredholm Resolvents, J. Comp. Syst. Sci. 2, 76-87 (1968).
- [11] Wiener N.: The Extrapolation, Interpolation and Smoothing of Stationary Time Series and Engineering Applications, Wiley, New York (1942).







**LT H**